METHOD, APPARATUS AND
COMPUTER PROGRAM PRODUCT
TO CRAWL A WEB SITE

5   ABSTRACT

In one embodiment, an improved method for crawling a web site is provided. At least

one page of the web site has a reference for executing by a browser to produce an address for a

next page. The web site is crawled by the crawler program, which includes querying the web site

server. The crawler parses such a reference from one of the web pages, and sends the reference

10   to an applet running in the browser. The address for the next page is determined by the browser

responsive to the reference. The address is then sent to the crawler. In an application of the

improved crawler, the crawler is used for reducing dynamic data generation on the web site

server. In this application, at least some of the web pages are dynamically generated responsive

to the crawler queries. The server generated web pages are processed to generate corresponding

15   processed versions of the web pages, so that the processed versions can be served in response to

future queries, reducing dynamic generation of web pages by the server.